

PadhAI: Variants of Gradient Descent

One Fourth Labs

Running and visualising momentum based gradient descent

Let's look at the Python implementation of Momentum based Gradient Descent

- Here is the Python code for Momentum Based Gradient Descent

```
X = [0.5, 2.5]
Y = [0.2, 0.9]

def f(w, b, x):
    # sigmoid with parameters w, b
    return 1.0 / (1.0 + np.exp(-(w*x + b)))

def error(w, b):
    err = 0.0
    for x, y in zip(X, Y):
        fx = f(w, b, x)
        err += 0.5 * (fx - y) ** 2
    return err

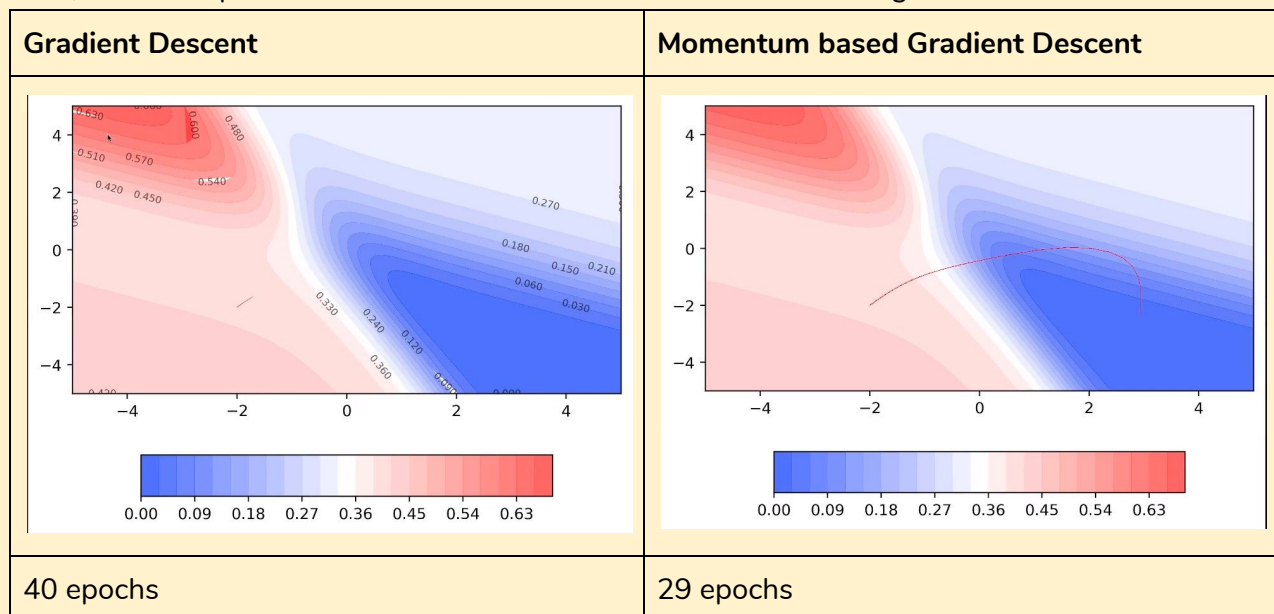
def grad_b(w, b, x, y):
    fx = f(w, b, x)
    return (fx - y) * fx * (1 - fx)

def grad_w(w, b, x, y):
    fx = f(w, b, x)
    return (fx - y) * fx * (1 - fx) * x

def do_momentum_gradient_descent():
    w, b, eta, max_epochs = -2, -2, 1.0, 1000
    v_w, v_b = 0, 0
    gamma = 0.7
    for i in range(max_epochs):
        dw, db = 0, 0
        for x, y in zip(X, Y):
            dw += grad_w(w, b, x, y)
            db += grad_b(w, b, x, y)
        v_w = gamma*v_w + eta*dw
        v_b = gamma*v_b + eta*db

        w = w - v_w
        b = b - v_b
```

- Now, let us compare the movement of Momentum based GD and regular GD



- However, there are still some issues with Momentum based GD that we will address in the next section